AD-A194 371

DOCUMENTATION PAGE

1b RESTRICTIVE MARKINGS	

2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Approved for public release; distribution is unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
R-140	

DTIC
SELECTED
JUN 0 2 1988
D

Accession For
NTIS CRA&I
DTIC TAB
Unannounced
Justification

By
Distribution/

Availability Codes

Dist | Avail and/or Special

A-1

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
USAETL	CEETL-LO	

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Fort Belvoir, VA 22060-5546	

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

DETERMINING VEHICLE MOTION FROM STEREO IMAGE SEQUENCES

12. PERSONAL AUTHOR(S)
ANN WERKHEISER

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
RESEARCH	FROM _____ TO _____	87 AUGUST 21	

16. SUPPLEMENTARY NOTATION

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	three-dimensional motion analysis
			parameters of motion
			symbolic matching system

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
Results are reported from a research project on three-dimensional motion analysis. The work is based on stereo image sequences of military vehicle. The analysis depends on the segmentation of the image and the matching of features between the two images. Preliminary results are given for an extension of the work to object recognition. Extracted features and the calculated parameters of motion are incorporated into a symbolic matching system. Object recognition is performed by the symbolic inexact matching of a representation of the photographs with a database of prototypes. The use of control points to furnish ground truth is discussed.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
F. DARLENE SEYLER	(202) 355-2647	CEETL-LO

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Determining vehicle motion from stereo image sequences

Anne Werkheiser

U. S. Army Engineer Topographic Laboratories, Research Institute
Bldg. 2592, Fort Belvoir, Virginia 22060-5546

ABSTRACT

Results are reported from a research project on three-dimensional motion analysis. The work is based on stereo image sequences of a military vehicle. The analysis depends on the segmentation of the image and the matching of features between the two images. Preliminary results are given for an extension of the work to object recognition. Extracted features and the calculated parameters of motion are incorporated into a symbolic matching system. Object recognition is performed by the symbolic inexact matching of a representation of the photographs with a database of prototypes. The use of control points to furnish ground truth is discussed.

1. INTRODUCTION

The work described is current research being performed collaboratively at the Center for Artificial Intelligence, U.S. Army Engineer Topographic Laboratories, and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign. Our long-term goal is to develop prototype components of a threat-detection system for a mobile robot. The research has centered on determining the three-dimensional motion of a vehicle in the field of view of the robot and, more recently, on attempting to recognize the vehicle.

2. APPROACH

We decided to work with sequences of stereo photographs of outdoor scenes because they present more realism and challenge than table-top models. Our motivation for using stereo is the instability of motion estimation algorithms based on monocular image sequences[1]. We included photogrammetric control so that we could have a form of ground truth to help evaluate success. Motion determination was first priority; parts of the motion work were subsequently incorporated into object recognition.

3. IMAGE DATA BASE

We first created a database of 18 stereo pairs by photographing an m114 armored personnel carrier (APC) in transit across the USAETL parking lot. Besides the APC, our pictures include trees, buildings, parked cars, a gas pump, and a

basketball hoop. We moved the APC to the first position, took one stereo pair, then moved it to the next position and stopped it for the next pair. The image sequence gives the effect of a vehicle in motion, but the motion is simulated. Everything but the vehicle is stationary, including the two cameras.

The imaging setup consisted of two Pentax 6 x 7 SLR cameras aligned (to the best of our ability) so that: the film planes lay in the same plane, the optical axes were parallel, and the plane containing them was parallel to the ground. The focal length of the lenses was 105 mm, the baseline was 3.6 m, and the distance from each optical axis to the ground was 1.22 m.

The distance from the vehicle to the cameras varies from 20 to 60 m. In addition, several stereo image pairs were taken with control points inserted in the scene. Each image (55 mm x 69.5 mm) was digitized on an Optronics drum scanner to 601 x 751 picture elements with 8 bits per picture element.

We subsequently photographed a second sequence of 20 stereo pairs with the same imaging setup but with two vehicles instead of one in some frames, not so much motion between frames as in the first sequence and a moving camera system.

We label our images Li, Ri which denote, respectively, the left and the right image of the ith stereo pair (i = 1, 2, ..., 18 or 20). The time instants at which the images were taken are

denoted by ti. For our experiments, sections of various sizes were windowed out of the images. Figure 1 shows a 128 x 256 pixel window of the left image in the ninth stereo pair of the first sequence.

Two consecutive stereo pairs at a time are processed, e.g. L1, R1, and L2, R2.

4. MOTION DETECTION AND ESTIMATION

We want to solve the following motion problem: Given two stereo image pairs taken at time instants t1 and t2 of a moving rigid object in a stationary natural surrounding, determine the 3-D motion and structure* of the object. Our approach is based on that described by Huang[2] in the section "Motion from 3-D Feature Correspondences." This method is:

Step 1. Detect and segment out the moving object in the images, thus eliminating the background.

Step 2. Extract and then match features (points, lines, circles) from the two images at t1 and then by triangulation determine the 3-D positions of these features. Do the same for the two images at t2.

Step 3. Match the two sets of 3-D points or other features at t1 and t2 to find 3-D point correspondences.

Step 4. Estimate the rotation and translation of the object from t1 to t2 by solving the set of equations involving the motion
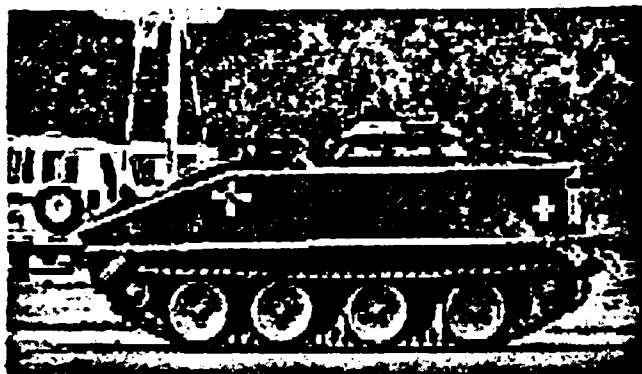

Figure 1. Left-9 image of m114 APC.

---

* "Structure" means the 3-D coordinates of the features that are extracted.

parameters.

We have an advantage over the monocular two-view case because, as Huang[2] notes, steps 2 and 3 are usually easier with stereo. For Step 2 this is because of the epipolar constraint; given a fixed point in one image of the stereo pair, the corresponding point in the other image is restricted to lie on the epipolar line. Step 3 has a rigid-body constraint; the distances between pairs of the 3-D points on a rigid body are invariant under motion.

The two constraints, epipolar and rigid body, give us an advantage, but there are two major difficulties.

The first difficulty is that in Step 2, it is hard to find feature extractors that yield common features in a stereo pair. Our experience with corner detectors, for example, is that the corners found in the right image very infrequently correspond with the corners found in the left image. In some experiments the set of corresponding corners comprised less that ten percent of the total. We got similar results using the Moravec[3] operator to detect "interesting" points. We were disappointed in both Burns lines[4] and Canny edges[5]. Very few of the Burns lines were common to the left and right images and Canny edges tended to contain parts of several objects. For example, an edge on a vehicle often merges into an edge on parked cars, trees, or other background objects.

The second difficulty is that in step 3, good results cannot be obtained. The problem is that range information on the features is usually very inaccurate[6] owing to image sampling.

To alleviate these difficulties, several new algorithms were developed. We finally achieved success by substituting for Steps 2 and 3:

Step 2a. Extract features in L1, and then find corresponding features in R1 by some form of cross correlation.

Step 2b. Similarly, extract and match features in images L2, R2.

Step 2c. Similarly, extract and match features in images L1, L2.

Step 2d. By taking the intersection of the three matched sets of 2a, 2b and 2c, obtain 4-way feature correspondences among images L1, R1, L2, R2.

Step 3. At t1 and t2, respectively, by triangulation determine the 3-D coordinates of the matched points.

Figure 2 shows the four-way matching. The best matching points came from a new algorithm based on zero-crossing patterns of second-order derivatives[7]. Figure 3 shows the eleventh left image from the second series with the edges extracted from zero crossings.

We are also encouraged by preliminary results with circle-finders[8]. Feature correspondences on circles can be used both for motion estimation and object recognition. Since the number of circular structures in a scene is normally small, finding correspondences is easier. Wheels usually present themselves as ellipses in imagery.
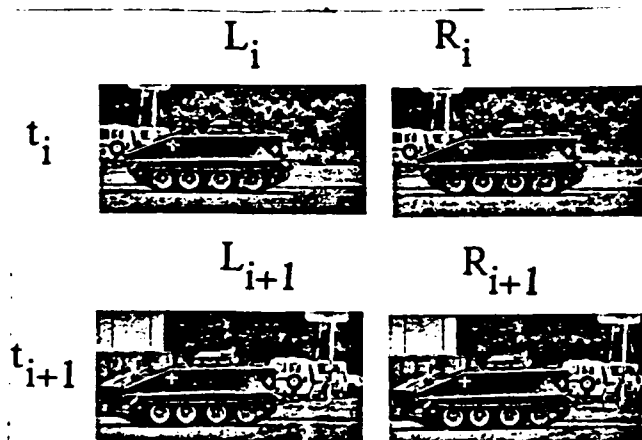

$$L_i \qquad R_i$$

t_i

$$L_{i+1} \qquad R_{i+1}$$

t_{i+1}

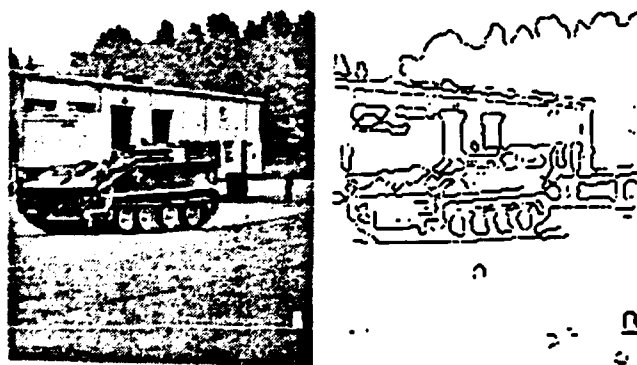Figure 2. Four-way point matching.

Figure 3. Edges from zero crossings.

McDonnell et al[8] discuss transforming the ellipses to circles.

In summary, after overcoming initial difficulties our project has developed reliable and robust algorithms for feature extraction[9], matching.[10,11] and solution of the equations involving the motion parameters[12]. Our experiments in using control points for ground truth have been reported[9]. We found our "ground truth" was good enough to show that our results are at least qualitatively correct. However a quantitative evaluation is not possible because our imaging setup was not calibrated carefully enough and the cameras were not metric.

5. OBJECT RECOGNITION

Horaud and Skordas[13] among others have observed that the task of recognizing an object in a random position and orientation is not trivial. An object in space has six degrees of freedom associated with it: three rotation parameters and three translation parameters. In our case the vehicle was constrained to move on a planar horizontal surface (the parking lot) so there was no translational motion up-and-down and the rotation is around an axis orthogonal to the ground plane. Only three degrees of freedom have to be determined, one rotation parameter and two translation parameters.

Our object recognition efforts began only after we had developed a reliable package for estimating motion parameters. We therefore assume we have the parameters and use them to help classify the vehicle. The motion parameters are the basis for our main heuristic: "The parameters of motion determine the orientation of the vehicle."

We use orientation as a means to index into the database of prototype vehicles. For example if the motion parameters show a large horizontal translation, we assume we are looking at the side of a vehicle, and we compare our photograph only with side views of prototypes.

Our database is composed of three dimensional Constructive Solid Geometry (CSG) models. A viewing angle for a model is specified as azimuth, α and an elevation. Our elevation is zero because the parking lot is a planar surface. The parameters of motion, a rotation matrix R and a translation vector T, are two dimensional. We

therefore have

$$\mathbf{R} = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} \quad \mathbf{T} = \begin{vmatrix} \Delta X \\ \Delta Y \end{vmatrix} .$$

The needed viewing angle, α, is computed as the sum of the two angles θ and $\phi = \tan^{-1} \Delta Y / \Delta X$.

Our paradigm for object recognition is the symbolic inexact matching of a candidate with several prototypes[14]. We use the BRL-CAD[15] solid modeling system to generate our prototypes. The package came with sample models of a tank and a truck, and we generated a model of the APC ourselves.

Figures 4, 5 and 6 show "wiresketches" of prototypes with hidden lines removed.

For the inexact matching itself, our intent was to use the work of Boyer, Vayda and Kak[16], which is a substantial extension of previous work by Shapiro and Haralick[14]. We found, however, that in our first preliminary experiments we were
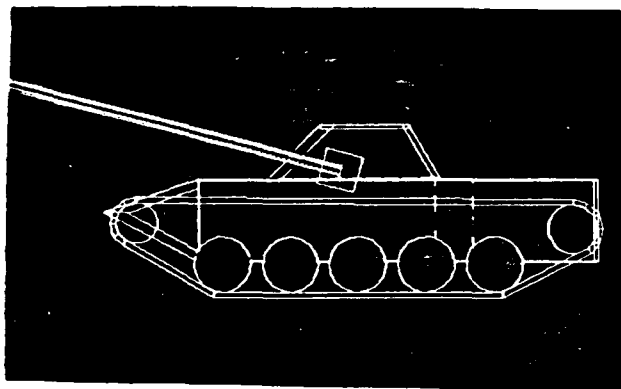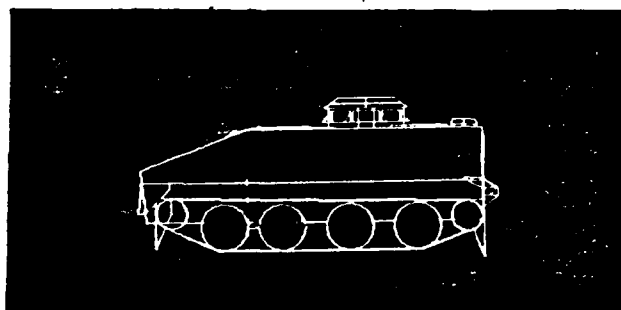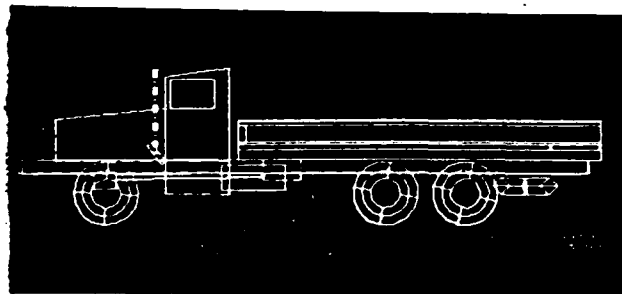

Figure 6. Wiresketch of a truck.

able to match our photograph to the correct wiresketch using the less elaborate method[11]. Part of the reason for this result could be that we made some simplifying assumptions (e.g. no noise features in the photograph) and that our partly-manual procedure did a better-than-average job of feature extraction. It also appears that our procedure gains strength from having two types of primitives to match, lines and circles, instead of only lines. We also have more constraints; for example some of our lines have to be parallel.

8. REFERENCES

1. J. Q. Fang and T. S. Huang, "Solving 3-D Small-Rotation Motion Equations", *Computer Vision, Graphics, and Image Processing*, Vol. 26 pp. 183-296, 1984.

2. T. S. Huang, "Motion Analysis" in *Encyclopedia of Artificial Intelligence*, Wiley, 1987.

3. H. P. Moravec, "Towards automatic visual obstacle avoidance", *Proc. 5th IJCAI*, August 1977, pp. 584-592.

4. J. B. Burns, A. R. Hanson, and E. M. Riseman, "Extracting Straight Lines", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 4, pp. 425-455, July 1986.

5. J. F. Canny, "Finding Edges and Lines in Images", *Tech Report 720, MIT AI Laboratory*, June, 1983.

Figure 4. Wiresketch of a tank.

Figure 5. Wiresketch of an armored personnel carrier.

6. S. D. Blostein and T. S. Huang, "Quantization Errors in Stereo Triangulation, to appear in IEEE Trans. Pattern Analysis and Machine Intelligence, 1988.

7. T. S. Huang, "Algorithms for Image and Scene Matching", *Final Report on Contract DAAL03-86-D-0001*, 29 June 1987. Engineer Topographic Laboratories, Fort Belvoir, VA 22060-5546.

8. M. M. McDonnell, M. Lew and T. S. Huang, "Finding Wheels of Vehicles in Stereo Images", this volume.

9. T. S. Huang, S. D. Blostein, A. Werkheiser, M. McDonnell, and M. Lew, "Motion Estimation from Stereo Image Sequences: Some Preliminary Results," *Proc. IEEE Workshop on Motion*, pp. 45-46, May 7-9 1986, Kiawah Island, Charleston, NC.

10. H. H. Chen and T. S. Huang, "Determining multiple object motion by 3-D point matching, to appear in *Pattern Recognition*, 1988.

11. K. S. Arun, T. S. Huang, and S. D. Blostein, "3-D Point Matching by Least Squares", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, September 1987, pp 698-700.

12. T. S. Huang, S. D. Blostein and E. A. Margerum, "Least Squares Estimation of Motion Parameters From 3-D Point Correspondences." Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, Fla., December 1986, pp 78-81.

13. R. Horaud and T. Skordas, "Model-Based Strategy Planning for Recognizing Partially Occluded Parts", *IEEE Computer Magazine*, Vol. 20, No. 8, August 1987, pp. 58-65.

14. L. G. Shapiro and R. M. Haralick, "Structural Descriptions and Inexact Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, 1981, pp. 504-519.

15. M. Muuss, P. Dykstra, K. Applin, G. Moss, E. Davisson, P. Stay and C. Kennedy, "Ballistic Research Laboratory CAD Package: A Solid Modeling System and Ray-Tracing Distribution Package." *SECAD/VLD Computing Consortium, The U. S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, 21005-5066.*

16. K. L. Boyer, A. J. Vayda, and A. C. Kak, "Robotic Experiments Using Structural Stereopsis for 3-D Vision." *IEEE Expert, Fall 1986*, pp. 73-94

END

DATE

FIlMED

8-88

DTIC